Find and Replace Missing Values in DataFrame

Aim

To write a Pandas program that finds and replaces missing values in a given DataFrame which do not have any valuable information.

Algorithm

1. Import the necessary libraries (pandas and numpy).
2. Create a sample DataFrame with some missing values.
3. Display the original DataFrame.
4. Replace the missing values:
    - For numeric columns, replace NaN with the mean of the column.
    - For string columns, replace NaN with 'Unknown'.
5. Display the DataFrame after replacing the missing values.

Code

```
import pandas as pd
import numpy as np

data = {
    'A': [1, 2, np.nan, 4, 5],
    'B': [5, np.nan, 7, 8, np.nan],
    'C': ['a', 'b', np.nan, 'd', 'e'],
    'D': [10, 20, 30, np.nan, 50]
}

df = pd.DataFrame(data)

print("Original DataFrame:")
print(df)

df['A'].fillna(df['A'].mean(), inplace=True)
df['B'].fillna(df['B'].mean(), inplace=True)
df['C'].fillna('Unknown', inplace=True)
df['D'].fillna(df['D'].mean(), inplace=True)

print("\nDataFrame after replacing missing values:")
print(df)
```

Output

```
Original DataFrame:
    A    B     C    D
0  1.0  5.0    a  10.0
1  2.0  NaN    b  20.0
```

```
2  NaN  7.0   NaN 30.0
3  4.0  8.0    d   NaN
4  5.0  NaN    e  50.0

DataFrame after replacing missing values:
      A         B        C        D
0  1.000000  5.000000    a      10.0
1  2.000000  6.666667    b      20.0
2  3.000000  7.000000  Unknown  30.0
3  4.000000  8.000000    d      27.5
4  5.000000  6.666667    e      50.0
```

Result

The program successfully found and replaced missing values in the DataFrame. Numeric missing values were replaced with the mean of their respective columns, while string missing values were replaced with 'Unknown'.